REMOTE SATELLITE POSITION & POSE ESTIMATION USING MONOCULAR VISION

Daniël F. Malan[*], **Willem H. Steyn** [#]

Key Words:
3D Tracking, Computer Vision, Kalman Filter, Structure from Motion, Satellite Formationkeeping

Abstract[**]

This thesis investigates methods for estimating relative 3D position and pose from monocular image sequences. The intended future application is of one satellite observing another, when flying in close formation. The ideas explored in this thesis build on methods developed for use in camera calibration and Kalman filter-based structure from motion (SfM). Each of the algorithms relies on visible feature points affixed to the target satellite with known geometry. To the author's knowledge, monocular vision in a Kalman filter milieu has not been previously used to estimate satellites' relative position and orientation. After describing the problem from a mathematical perspective we develop different approaches to solving the estimation problem. The different approaches are successfully tested on simulated as well as real-world image sequences, and their performance analyzed. Results show the algorithms to be successful in tracking simulated as well as physical targets. The effectiveness of a direct least-squares solution versus a stand-alone Extended Kalman Filter (EKF) or Unscented Kalman Filter (UKF) is investigated. The recently developed Unscented Kalman Filter is found to be less suited to our application than the more widely used EKF.

1. INTRODUCTION

A**lthough** satellite constellations are in widespread use, the application of satellites flying in close formation (separated by less than 200 m) is relatively new. The University of Stellenbosch's "Electronic Systems Laboratory" is currently looking at using close formation flying as part of future micro-satellite missions. The use of close formation flying between imaging satellites can be used for stereo earth imaging, and is also useful for external observation of a "mother" satellite by a "slave" satellite.

Accurate information concerning relative position and motion is required to maintain formation flight between satellites, and might be useful for maneuvers such as satellite docking. Information concerning relative orientation might be important when doing inspection, for alignment of antennae, and also for docking. These requirements stipulate the need for accurate estimation of relative position and pose. Optical computer vision could represent an affordable and practical way to determine position and orientation between an observer and a target.

---

[*] D.F. Malan is an M.Sc student at the Department of Electrical and Electronic Engineering, University of Stellenbosch, South Africa. Phone: +27 21 808 4472, email: fmalan@sun.ac.za

[#] W.H. Steyn is Professor in Satellite and Control Systems at the Department of Electrical and Electronic Engineering, University of Stellenbosch, email: whsteyn@sun.ac.za

[**] The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the NRF.

We propose that the position and pose estimates be computed locally by one of the satellites (the observation satellite). This approach is desirable since it circumvents any transmission delays, and is not dependent on base station observability. The onboard processing power of a micro-satellite should be sufficient for required task.

Research focusing on motion estimation from monocular image sequences over the past fifteen years often addressed the problem of structure from motion (SfM). SfM attempts to estimate the 3D motion as well as the 3D structure of a target, by analyzing its 2D image (video) projection. A problem with SfM is that its estimates are typically scale invariant. The shape of the target can be determined, but the unknown size of the target results in an undetermined distance from the observer. This problem is unique to monocular estimation, where triangulation is impossible. In our application we are explicitly interested in the absolute 3D position (and pose) of the target. Furthermore, the dimensions and structure of a satellite target will be known to the observer. By incorporating the knowledge of the target being viewed, we can attempt to apply the techniques of SfM to estimating the absolute translational and rotational state of the target. Since we are interested in determining a subset of SfM's unknowns the problem should be more tractable.

Two fundamentally distinct approaches exist for obtaining 2D measurement data for motion estimation. Optical flow based methods represent 2D motion in the image plane as a continuous velocity field. This approach has proved useful in applications where a complex scene needs to be segmented into moving and stationary components. Feature based methods, on the other hand, rely on the recognition of a set of corresponding feature points as they occur in consecutive image frames. We will be looking at feature-based methods.

A seminal paper by Broida, Chandrashekhar and Chellappa [1] focused on using corresponding feature points between successive images to estimate both 3D motion and structure. Measurements of 3D accuracy were not obtained due to scale invariance of the structure estimator, and lack of "ground truth" measurements for the real-world motion sequences.

An attempt at absolute pose and motion estimation by using line features was investigated by Goddard [2]. In this case absolute 3D accuracy was measured, but only for uniform motion with the target placed close (≤ 1 meter) to the observer.

Attempts at SfM estimation by using automatically extracted feature points and UKF filtering was investigated by Venter [3], and is currently being developed further by Rautenbach [4]. All of these investigations use various nonlinear extensions of the Kalman Filter as a central element of the estimation algorithms.

There are plausible alternative ways in which relative 3D position could be determined between satellites. Stereo vision (as is human vision) has the advantage of enabling 3D triangulation which is especially well suited to identifying the structure of an unknown object. The relative distances, the fact that we track a known target, and space and weight limitations count against such an approach. Alternatively relative position and pose can be computed by transmitting one satellite's onboard orientation estimate to the other. A third alternative is measuring range with radar or laser rangefinder, although this provides no attitude information.

2. METHODOLOGY

We propose the use of a single digital camera (monocular vision) to track the target (satellite). The digital camera's detector might, for example, consist of a 1 megapixel monochrome CMOS or CCD imager. Visible markers are affixed at known positions on the target to be used as point features.

We modeled the orbital kinematics of a satellite, as well as the 3D object to 2D image plane projection process. A perspective (pinhole) camera model was used.

Each observation (image frame) will provide us with a certain number of visible feature points. The first step in the tracking algorithms would be to identify the feature points in the 2D image frame. The second step involves identifying each feature and matching it with its known position on the target. The third (and most important) step is to incorporate these feature matches to an estimate of position and pose. We can use an algorithm to combine all features of a given image frame to a single "measurement" of position and pose, or we can treat each point feature as a separate measurement - thereby avoiding the use of algorithms which depend on the number of visible features. Both approaches are explored.

Different variations of the Kalman Filter (KF) are implemented to recursively estimate the 3D position and pose of the target. The process of inverting the 3D to 2D imaging process can be solved by minimizing a set op linear equations in a least squares sense. The output from this so-called "calibration algorithm" can be smoothed by applying a KF. The KF, however, is a potentially powerful tool, which can in itself invert the imaging process by means of its internal measurement model. We compare methods relying either on a calibration / KF combination, with methods using only a nonlinear KF. We examine the performance of the well-known Extended Kalman Filter (EKF) when compared to the recent and much vaunted Unscented Kalman Filter (UKF). We furthermore investigate the sensitivity of each of the methods to measurement noise and number of visible features.

The algorithms under scrutiny are tested on simulated as well as real image sequences. All algorithms and simulations were implemented and tested in the Matlab environment. The real-world image sequences were captured by using a standard consumer digital camera, and processed offline. The simulated "image sequence" was chosen to represent a scenario which with believable values for the intended satellite application.

3. CONCLUSIONS

After running tests on both the simulated and the real-world (see Figure 1) image sequences we found that the so-called calibration algorithms are more noise-sensitive than purely Kalman-filtered trackers, and need at least six visible point correspondences to generate reliable estimates (which was not a problem for our experiments). Due to the fact that a calibration algorithm's least-squares solution does not rely on any previous estimate, it can be useful for initializing the position and pose state vector, for Kalman filtering.

It was found, however, that using purely a full nonlinear Kalman Filter (such as either an EKF or UKF) provided much better tracking and tolerance to image noise than the aforementioned calibration algorithms, of which the output is smoothed. Similar to Broida [1] we found that convergence of the nonlinear EKF or UKF filter can be a problem under conditions of high image noise or bad initial estimates. This could mostly be remedied by careful choice of the filter's covariance matrices.

An interesting result is that the UKF did not provide better estimates than the EKF in our tests. This supports some experimental work by La Viola [6] for quaternion estimation. When taking the increased computational complexity of the UKF into account it is not recommended for this application. We do however agree that the UKF might prove more flexible for very complex nonlinear applications (such as full SfM), although this falls outside the scope of this investigation.

Results from one of the practical tests are shown beneath. The target was placed three meters in front of the camera, and the images captured at 640x480 pixels at 30fps.

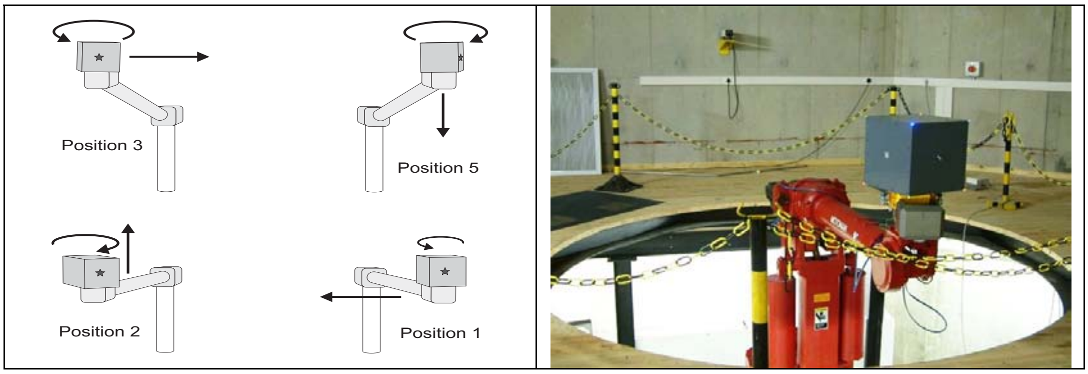

Figure 1: Target fitted with LED markers being manipulated by a robot arm

Method used	**RMS Tracking error**			
	Position (m)	Velocity (m/s)	Rotation (degrees)	Angular rate (deg/s)
Cal & LKF+EKF	0.065	0.033	3.8	3.2
Cal & LKF+UKF	0.065	0.033	4.3	8.3
EKF	0.057	0.027	3.4	2.7
UKF	0.064	0.072	6.0	9.9

Table 1: Tracking accuracy for the target of Figure 1

REFERENCES:

[1] T.J. Broida, S. Chandrashekhar, and R. Chellappa, “Recursive 3-d motion estimation from a monocular image sequence”, IEEE Transactions on Aerospace and Electronic Systems, vol. 26, no. 4, pp. 639--656, July (1990)
[2] James S. Goddard Jr., “Pose and Motion Estimation from Vision using Dual Quaternion-based Extended Kalman Filtering”, Ph.D. thesis, The University of Tennessee at Knoxville, (1997)
[3] Chris Venter, “Structure from motion estimation using a nonlinear Kalman filter”, M.Sc. thesis, University of Stellenbosch, (2002)
[4] Pieter Rautenbach, “3d facial reconstruction by means of structure from motion", M.Sc. thesis, University of Stellenbosch, (2004)
[5] Joaquim Salvi, Xavier Armangué, and Joan Batlle, “A comparative review of camera calibrating methods with accuracy evaluation”, Pattern Recognition, vol. 35, pp. 1617-1635, (2002)
[6] Joseph J. La Viola, “A comparison of unscented and extended Kalman filtering for estimating quaternion motion”, Proceedings 2003 - American Control Conference, pp. 2435--2440, (2003)